

Contribuciones cortas

Propuesta de términos para la indización en ciencias de la información

Soledad Díaz del Campo

Resumen

Como parte de las acciones emprendidas para responder a la necesidad de superación y de actualización permanente de los especialistas y técnicos del Sistema Nacional de Información de Ciencias Médicas, se diseñó un vocabulario para la indización de los artículos que se publican en la revista ACIMED, así como de los reprints que conforman la base de datos RACI. En el presente artículo se dan a conocer, en apretada síntesis, las fuentes de datos utilizadas y la metodología adoptada para confeccionar el vocabulario y se detallan otros aspectos en relación con su estructura, la forma de presentación de sus términos y sus aplicaciones.

Descriptores (DeCS): VOCABULARIO; RESUMEN E INDIZACION; CIENCIAS DE LA INFORMACION

Descriptores (DeCI): VOCABULARIO; INDIZACIÓN; CIENCIAS DE LA INFORMACIÓN

Abstract

As part of the actions carried out to respond to the need of constant upgrading and up-dating of specialists and technicians in the National Medical Science Information System, a vocabulary stock was designed for the indexing of the articles published in the journal ACIMED, as well as the reprints included in the RACI data base. In this article information is provided, in a synthetic manner, on the data sources used and the methodology applied to make up the vocabulary, and other aspects related to its structure, term presentation and applications are detailed.

Subject headings (DeCS): VOCABULARY; ABSTRACTING AND INDEXING; INFORMATION SCIENCE

Subject headings (DeCI): VOCABULARY; INDEXING; INFORMATION SCIENCES

El hecho de que las lenguas más utilizadas a través del tiempo en las distintas disciplinas científicas hayan rebasado los límites de sus respectivas naciones de origen, se explica en el carácter internacional de las ciencias. En un principio, fue el griego el idioma predominante en toda Europa. Luego, con el advenimiento de la Edad Media, este cedió su lugar al latín, mientras que durante los siglos XVII y XVIII fue el alemán el que sobresalió como idioma del sistema de conocimientos sobre la naturaleza, la sociedad y el pensamiento.

Por otra parte, conviene recordar que por aquella época se tenían por sabios a quienes guardaban en su

memoria un enorme bagaje de información. Estos individuos eran como compendios vivos de las reglas de las ciencias y las artes, capaces de reemplazar entonces, a los verdaderos prontuarios. Esa posibilidad existía porque el volumen de conocimientos no era en realidad muy grande. Un ejemplo elocuente es el del célebre filósofo y matemático alemán Gottfried Wilhelm Leibnitz (1646-1716) quien, con la gran cantidad de trabajos que redactó y publicó en latín y en francés y, por supuesto en alemán, demostró la capacidad de los científicos de aquellos tiempos de comunicarse en varias lenguas.¹

En el siglo XIX comenzó a vislumbrarse el cambio diametral que se produciría luego en la hegemonía lingüística, por cuanto desde sus inicios se produjo un auge de la influencia del inglés, el cual llegó a convertirse en la "lengua franca" de las ciencias, sobre todo tras la Segunda Guerra Mundial, cuando logró desplazar al alemán en el contexto de las ciencias naturales. Este hecho trajo consigo que los propios investigadores germanoparlantes suelen publicar desde entonces en ese idioma sus resultados más importantes.²

Cuando se habla en específico de las ciencias de la información, se hace referencia a un aspecto importante del acontecer científico y cultural de la época actual, el cual más que internacional, es ya indispensable en cualquier esfera de la vida social. Así, las ciencias de la información forman parte de todo lo que tiene que ver hoy día con la evolución y el perfeccionamiento de la especie humana y de su entorno, con independencia de que su valiosa contribución al logro de una mejor visión del mundo ha conllevado la necesidad de buscar recursos que permitan al menos asimilar la mayor cantidad de sus potencialidades en función de un fenómeno concreto. Ya quedaron atrás los tiempos en los que un individuo o un reducido grupo de personas era capaz de asimilar la mayor parte o una gran proporción de los conocimientos disponibles.

Entre los recursos de ayuda a la recuperación efectiva de la información ante el desafío impuesto por la actual revolución en el proceso de comunicación para poder acceder al conocimiento, sobresalen los lenguajes de búsqueda informativa con aptitudes para asimilar los atributos del lenguaje natural, tanto desde el punto de vista lógico como en el aspecto semántico, ya se trate del inglés, el español, el alemán, etc.

Si bien la existencia de este tipo de vocabulario data de hace más de una centuria, la asignación de epígrafes o de descriptores cobra actualmente un valor incalculable, toda vez que se vive una época de grandes transformaciones, no sólo en las formas de conocer y de comunicarse, sino también en los modos de hacer y de pensar en cualquier actividad de la especie humana.^{3,4} De ahí su carácter de vehículos facilitadores de inteligencia y, por tanto, de elementos importantes para la productividad científica.

De lo anterior se infiere que los también llamados "lenguajes artificiales", tradicionales, indicadores del contenido de los documentos o de las solicitudes de información, tienen hoy día aplicación en cualquier esfera del conocimiento, con inclusión de las ciencias de la información.

El objetivo del presente trabajo es dar a conocer de modo sintético los pormenores que condujeron a la creación de un vocabulario con términos para la indización en ciencias de la información, concebido para responder en principio a la necesidad de superación y actualización constante del personal que labora en el Sistema Nacional de Información de Ciencias Médicas (SNICM). Asimismo se indica la estructura de dicho vocabulario, la forma de presentación de sus términos y sus diversas aplicaciones.

Métodos

Los especialistas y técnicos del SNICM están familiarizados con los tesauros Medical Subject Headings (MeSH) y Descriptores en Ciencias de la Salud (DeCS), los cuales son dentro de ese medio los máximos exponentes del proceso de transformación lingüística para representar y describir contenidos documentarios y solicitudes de información sobre las ciencias de la salud y sus disciplinas afines.^{5,6}

Estos profesionales, además de facilitadores de conocimientos, cuyos aportes como tales se refleja en el perfeccionamiento cualitativo de la asistencia, la investigación y la docencia médica, son a su vez consumidores y generadores de información dentro de la rama que cultivan, razón por la cual deben disponer de mecanismos que faciliten potenciar también su quehacer en este sentido. Para responder a ese requerimiento, se ha puesto en práctica más de una iniciativa con positivos resultados,^{7,8} sin embargo, el propio desarrollo de estos proyectos ha conllevado nuevas exigencias, a las cuales ha sido necesario hacerles también frente. Por ejemplo, el hecho de agregarle valor a la producción de la revista ACIMED con

la traducción de los resúmenes al inglés y la indización de los artículos con vistas a su inclusión en bases de datos internacionales,⁹ significó un nuevo desafío a partir de 1996, al igual que el empeño de crear y poner en explotación una base de datos para beneficio de los mismos profesionales.⁸

Para encarar este reto, se comenzaron a indizar los artículos de ACIMED y los reprints de la base de datos RACI con el vocabulario DeCS, formado por categorías, en las cuales se agrupan los descriptores afines a determinada materia. Así, la categoría A para Anatomía, la B para Organismos, la C para Enfermedades, la L para las Ciencias de la Información, etc.¹⁰

Desde el mismo comienzo de este trabajo se detectó la insuficiencia de los descriptores de la categoría L del DeCS para la indización de todos los aspectos que abarcan actualmente las ciencias de la información. Este inconveniente dio lugar a la búsqueda de una solución inmediata, dada la necesidad imperiosa de contar con un vocabulario que pudiera reflejar de modo permanente el contenido de los artículos publicados en ACIMED y de los reprints que alimentan la base de datos RACI. Al efecto se ejecutaron varias acciones en diferentes etapas, la primera de las cuales consistió en visitar algunas bibliotecas con el fin de conocer los vocabularios que utilizan para la indización. Como resultado de esta gestión se supo que dicha labor se hace con palabras clave sobre las cuales no se tiene control alguno.

El punto de partida de la segunda etapa fue la consulta de tesauros, diccionarios, trabajos de diploma y otros documentos vinculados con el tema, de los cuales se extrajeron términos que luego se ordenaron alfabéticamente en un fichero con su equivalente en inglés y se sometieron a una minuciosa revisión para evitar sinonimia. Asimismo procede informar que desde el comienzo de los trabajos de indización con los 700 términos seleccionados de inicio, se notó la necesidad de disponer de listados auxiliares y de ampliar el número de calificadores adscriptos en el DeCS a la categoría L. Esto condujo a dedicar la tercera etapa a la agrupación de los términos obtenidos según áreas temáticas, para formar cadenas jerárquicas con su correspondiente encabezamiento en cada caso. Con posterioridad se introdujeron todos los términos en el programa DECSA, soportado en CD-ISIS, versión 3.7 de 1993, para lograr el índice permutado y se emprendió el proceso de ordenamiento alfabético de los términos en idioma inglés con su equivalente en español. Esta última acción obedece a que el 90% de los reprints que conforman la base de datos RACI están originalmente publicados en ese idioma.

Por otra parte, la necesidad de ampliar el número de calificadores tuvo su origen en que los 14 adscriptos a la categoría L del DeCS fueron insuficientes para expresar con el nivel de detalle requerido el contenido de una gran cantidad de reprints y de los artículos de temáticas tan diversas publicados en la revista ACIMED. Entonces se analizó el número de veces que fue necesario usar otros calificadores para luego adoptarlos (anexo 1).

La cuarta etapa abarcó la incorporación de términos tomados de los textos de los propios reprints y la proyección de otros elementos necesarios para la conformación final del vocabulario, tales como la introducción, la forma de presentación de los términos, las abreviaturas utilizadas, los calificadores, etc. El trabajo de composición se llevó a cabo en lenguaje Microsoft Word de Window 98.

Estructura

Este vocabulario, cuyo título responde a su cualidad de ser una "Propuesta de Términos para la Indización en Ciencias de la Información", consta de cinco páginas preliminares, donde aparecen la portada, la tabla de contenido, la introducción, los agradecimientos, la forma de presentación de los términos, las abreviaturas y los calificadores. A estas les siguen otras 187 páginas dedicadas a la lista alfabética, compuesta por 1 225 términos; la lista jerárquica, la lista permutada y la lista de términos en inglés.

Lista alfabética

Como su nombre lo indica, en ella aparecen los términos en orden alfabético con su equivalente en inglés. Algunos de estos vocablos van acompañados de notas de alcance para la mejor orientación del indizador. A la mayoría de los términos se adjuntan también el encabezamiento de la cadena jerárquica a la que pertenece y su correspondiente término genérico; así como los términos específicos y relacionados en los casos que los tuvieran. Esta lista es la principal fuente para la indización (anexo 2).

Lista jerárquica

En esta parte se agrupan los términos en cadenas jerárquicas. De ahí que a continuación de un vocablo que sirve de encabezamiento, se ubiquen otros términos a él subordinados. Su utilidad radica en que permite indizar con un mayor grado de especificidad (anexo 3).

Lista permutada

Aquí están consignadas alfabéticamente todas y cada una de las palabras que componen los términos simples o compuestos del vocabulario. Su utilidad está dada en que por medio de ella se puede lograr una indización más exacta a partir del conocimiento de una sola de las palabras integrantes de cualquier término (anexo 4).

Lista de términos en inglés

En esta relación se encuentran los mismos términos de la lista alfabética, expresados en idioma inglés con su equivalente en español. Su existencia obedece a la eventual necesidad de procesar documentos originalmente redactados en ese idioma (anexo 5).

Consideraciones generales

Dotar al trabajador de la información de una herramienta para que este lleve a cabo en su propia disciplina la labor de indización de una manera más eficaz, es una exigencia igual a la que se hizo sentir al momento de surgir la de que cada ciencia contara con un conjunto de vocablos normalizados o, al menos, reconocidos por la comunidad científica; del mismo modo que la evolución y desarrollo del lenguaje hablado conllevó la urgencia de crear soportes donde se registraran, en un orden determinado, las voces y giros idiomáticos para hacer más transparente la capacidad de comunicación intrínseca del hombre.

Como se puede inferir de la lectura de este artículo, los términos que se proponen para la indización en ciencias de la información tratan de dar una modesta respuesta a esta cada vez mayor exigencia. Sin pretender en principio que con ellos los adeptos a la indización dispongan ya de un instrumento acabado para abrirse camino en la complicada urdimbre de esta actividad, el solo hecho de poder divulgar a través de estas páginas la existencia del vocabulario posibilita, su perfeccionamiento con las valiosas sugerencias de los interesados en el tema.

Vale recalcar que este vocabulario, más que una mera serie de palabras reunidas por obra de la imaginación, es el resultado de una engorrosa y dilatada consagración al estudio y a la coordinación de los términos más usados en una ciencia polimorfa que evoluciona de modo vertiginoso y que, en consecuencia, trae consigo la aparición al vuelo de gran cantidad de nuevas expresiones. Procede también subrayar que el plan seguido en la experimentación es rigurosamente científico y original, aun cuando guarden alguna semejanza con trabajos de otros autores que hayan abordado el asunto.

La obra aquí descrita de manera tan somera se emprendió porque se creyó útil, dada su posibilidad de llenar un vacío. Por eso este primer esfuerzo será doblemente honrado, si en verdad resulta provechosa a los indizadores, y si puede por lo menos servir de antecedente a otra que se acerque más a la perfección.

ANEXO 1 Calificadores

<i>análisis</i>	<i>anal</i>	<i>AN</i>
(analysis)		
<i>aptitud</i>	<i>aptit</i>	<i>AL</i>
(ability) (abil)		
<i>*clasificación</i>	<i>clas</i>	<i>CL</i>
(classification) (class)		
<i>control</i>	<i>contr</i>	<i>CT</i>
(control)		
<i>desventajas</i>	<i>desv</i>	<i>DE</i>

(disadvantages) (disav)	
durabilidad.....durab.....	DU
(durability)	
*economía	econ EC
(economics)	
*educación	educ ED
(education)	
efectividad.....	effect EF
(effectiveness) (effect)	
eficiencia.....	efic..... EY
(efficiency) (effic)	
*estadísticas y datos numéricos	estadist SN
(statistics and numerical data) (statist)	
fiabilidad.....	fiab RY
(reliability) (relia)	
habilidades.....	habil..... SK
(skilfulness) (skilf)	
*historia	hist HI
(history)	
*instrumentación	instrum IS
(instrumentation)	
*legislación y jurisprudencia	legis LJ
(legislation and jurisprudence)	
*métodos	métodos MT
(methods) (methods)	
*normas	normas ST
(standards) (stand)	
*organización y administración	org OG
(organization and administration) (organ)	
perspectivas.....	pers..... PE
(perspective)	
*provisión y distribución	provis SD
(supply and distribution) (supply)	
*recursos humanos	rec hum MA
(manpower) (man)	
*tendencias	tend TD
(trends) (trends)	
*utilización	util UT
(utilization)	
ventajas.....	vent..... AV
(advantages) (adv)	

Leyenda:

*** Calificadores DeCS**

ANEXO 2 Lista alfabética

[illegible][illegible]

ABREVIATURAS ABBREVIATIONS	
ACADEMICAS DISERTACIONES ACADEMICAS	ADQUISICION ACQUISITION OF KNOWLEDGE
ACCESIBILIDAD ACCESSIBILITY	ADQUISICIONES ACQUISITIONS
ACCESO ACCESO A LA EDUCACION ACCESO A LA INFORMACION ACCESO A LOS MERCADOS ACCESO AL TRABAJO	CATALOGOS DE ADQUISICIONES
ACTAS ACTS	AFGANISTAN AFGHANISTAN
ACTIVIDADES ACTIVIDADES CIENTIFICAS INFORMES DE ACTIVIDADES	AFRICA AFRICA AFRICA AUSTRAL AFRICA CENTRAL AFRICA DEL NORTE AFRICA OCCIDENTAL AFRICA ORIENTAL AFRICA DEL SUR DEL SAHARA
ACUERDOS ACUERDOS ECONOMICOS ACUERDOS INTERNACIONALES	AFRICANA REPUBLICA CENTRO AFRICANA
ADecuACIONES ADECUACIONES	AGRADECIMIENTOS AGRADECIMIENTOS
ADMINISTRACION ADMINISTRACION ADMINISTRACION DE BIBLIOTECAS ADMINISTRACION DE LA DEMANDA ADMINISTRACION DE LA INFORMACION ADMINISTRACION DE LA OPERTA SISTEMAS DE ADMINISTRACION DE BASES DE DATOS	ALBANIA ALBANIA
ADMINISTRATIVA SYSTEMS DE INFORMACION ADMINISTRATIVA	ALEMANIA GERMANY
	ALGORITMOS ALGORITHMS
	ALMACENAMIENTO ALMACENAMIENTO DE DATOS ALMACENAMIENTO DE LA INFORMACION ALMACENAMIENTO MAGNETICO EQUIPO DE ALMACENAMIENTO DE COMPUTADORA

ANEXO 5 Lista de términos en inglés

ABBREVIATIONS ABREVIATURAS	AFGHANISTAN AFGHANISTAN
ABSTRACTS RESUMENES	AFRICA AFRICA
ACCESS TO EDUCATION ACCESO A LA EDUCACION	AFRICA, CENTRAL AFRICA CENTRAL
ACCESS TO INFORMATION ACCESO A LA INFORMACION	AFRICA, EASTERN AFRICA ORIENTAL
ACCESS TO MARKET ACCESO A LOS MERCADOS	AFRICA, NORTHERN AFRICA DEL NORTE
ACCESS TO WORK ACCESO AL TRABAJO	AFRICA SOUTH OF THE SAHARA AFRICA AL SUR DEL SAHARA
ACCESSIBILITY ACCESIBILIDAD	AFRICA, SOUTHERN AFRICA AUSTRAL
ACCESSIONS LIST CATALOGOS DE ADQUISICIONES	
ACCOUNTING CONTABILIDAD	
ACKNOWLEDGMENTS AGRADECIMIENTOS	
ADDED VALUE VALOR AÑADIDO	
ADQUISITIONS ADQUISICIONES	
ADVERTISING PUBLICIDAD	

Referencias bibliográficas

1. Asimov I. Gottfried Wilhelm Leibnitz. En: Asimov's biographical encyclopedia of science and technology. The lives and achievement of 1510 great scientifics from ancient times to the present chronologically arranged. 2 ed rev. Garden City: Doubleday, 1982:155-6.
2. Oksaar E. Deutsch als Wissenschaftssprache. Geburtsh u Frauenheilk 1997;57(6): M106-8.

3. Melamut AL, Otero PD. Acerca de la red de redes Internet. Arch Arg Pediatr 1997; 95(5):366-71.
4. Echenique Elizondo M. Internet y cirugía hoy. Arch Fac Med Zaragoza 1998;38(1):44-8.
5. National Library of Medicine. Medical Subject Headings: alphabetic list. Bethesda: US Department of Health and Human Services, 1996.
6. BIREME. Descriptores en Ciencias de la Salud. (DeCS): lista alfabética. Sao Paulo: OPS, 1992. (Metodología LILACS; 4.1)
7. López Espinosa JA. Editorial. ACIMED 1993;1(1):3.
8. Díaz del Campo S, López Espinosa JA. RACI: una base de datos para el profesional de la información. ACIMED 1999;7(1):40-4.
9. Hernández Valdés E. Hacia el primer lustro de ACIMED [editorial]. ACIMED 1997; 5(1):5-7.
10. BIREME. Descriptores en Ciencias de la Salud (CeCS): lista jerárquica. Sao Paulo: OPS, 1992. (Metodología LILACS; 4.2).

Recibido: 2 de diciembre del 2002

Aprobado: 18 de enero del 2003

Tec. Soledad Díaz del Campo

Editorial Ciencias Médicas (ECIMED).

Calle I No. 122 esq. Línea, piso 10, El Vedado, Ciudad de La Habana. Cuba.

1Técnica en Bibliotecología Médica. Editorial Ciencias Médicas (ECIMED).

© 2004 2000, Editorial Ciencias Médicas

Calle E No. 452 e/ 19 y 21, El Vedado, La Habana, 10400, Cuba.



acimed@infomed.sld.cu